DISCUSSION TOPICS WITH NMED
THURSDAY JUNE 24, 2004

- Laboratory analytical results for DMP well WQSP-6 have arrived and the data have been reviewed and validated. CBFO will provide these data to NMED as requested.

- In reference to the NMED comment about averaging duplicates samples in establishing the background baseline, the EPA guidance referenced by NMED does not require or recommend averaging duplicate samples in the establishment of a background baseline.

- The baseline and the time trend control charts show all data, both sample and duplicate concentration values. Using all data, not averages, provides a larger more representative sample of the population and demonstrates the inherent variability between sampling rounds and between single round laboratory analyses (precision) of WIPP brine groundwaters.

- Averaging sample and duplicate analyses when establishing background offers no advantage over incorporating all of the individual analytical results. Using all data provides greater confidence in the calculated statistics. Another advantage of the independent sample approach is that the minimum and maximum values and the standard deviation of the resulting distribution will be more representative of the true population range.

- In reference to the NMED request for the creation of a new spreadsheet with all of the analytical results for each well, the semiannual Detection Monitoring Reports that are submitted to NMED contain all of the requested data, for each well and each parameter, in spreadsheet format. NMED has the information requested here, including the Round 17 results. CBFO can provide these spreadsheets to NMED electronically, if desired.

- In reference to the NMED comment concerning analytical values that were rejected, CBFO will provide additional information and clarification relative to the decision to reject the few selected data points.

040630

Shaw The Shaw Group Inc.

Shaw Environmental & Infrastructure, Inc.

5301 Central Avenue NE
Albuquerque, NM 87108-1513
505-262-8600
Fax: 505-262-8855

Memorandum

Date: June 18, 2004

To: Mark Crawley

From: Dr. Jonathan Myers

RE: Response to NMED Comments on Averaging Duplicates

This memo provides a response to the NMED's objection to the handling of duplicate samples obtained from the groundwater monitoring program.

Regulatory Guidance

EPA, 1989 and 1992 guidance does not mention handling of duplicate sample results, except for a section in EPA, 1992 on the preparation of Combined Shewhart-Cumulative Sum Control Charts (CUSUM), which is one particular type of control chart that is not applicable to WIPP groundwater monitoring data. There are many different types of control charts that are applicable to different types of data. The CUSUM chart described in EPA, 1989 is only applicable for normally distributed data (EPA, 1989 emphasis), which does not include the left-censored, nonparametric distributions that characterize the trace elements in the WIPP data set.

In a CUSUM chart, the first several rounds of baseline monitoring data are used to calculate a mean and standard deviation. Subsequent data obtained during the monitoring phase are converted to a normalized z-value (distance above or below the baseline mean in units of standard deviation), and are plotted on the y-axis as a cumulative sum as a function of time (sample date) which is shown on the x-axis. This approach is obviously only applicable for normally distributed data for which an unbiased baseline mean and standard deviation can be established. The high percentage of nondetects in the RCRA-listed metals data precludes the calculation of an accurate mean or standard deviation, nor can the normality of the distribution even be tested. Guidance in EPA, 1989 states that distributions that fail tests for normality or lognormality, as well as any distributions with greater than 15 percent nondetects, be treated nonparametrically. For these reasons, we have been showing the data on nonparametric control charts. This approach does not require the calculation of a mean and standard deviation, and is applicable to data with any type of distribution and any percentage of nondetects.

Statistical Approach

There are advantages and disadvantages in averaging duplicate samples. A decision on whether to average duplicate samples prior to calculation of summary statistics should be based on maximizing the amount of information that can be extracted from the data set, while avoiding any bias that may be introduced by treating each analysis of each sample as an independent value. In the case of WIPP, summary statistics are separately calculated for each well. If each sample is treated independently, then sample rounds with duplicate samples will have greater weight in the summary statistics relative to rounds that only have analyses of single samples. However, no summary statistics are calculated for the monitoring phase data. Individual concentrations are simply compared to the baseline data. The only key statistical parameter that was calculated for the baseline data was the UTLV. If there is a seasonal effect on analyte concentrations, then a high or low bias may be introduced in the UTLV if duplicate samples happen to be obtained during the seasons when concentrations are high or low, respectively. However, the depths to groundwater at the hydrologic units that are sampled at the WIPP are such that the units are protected from seasonal fluctuations, so no such bias is introduced by treating each sample as an independent value.

The advantage of the independent sample approach is that the summary statistics are directly based on a larger sample size, so greater confidence can be placed in the calculated statistics. For instance, one has more confidence in a mean calculated from 20 samples relative to a mean calculated from 10 samples, even if the two resulting means are the same. Another advantage of the independent sample approach is that the minimum and maximum values and the standard deviation of the resulting distribution will be more representative of the true population range. Averaging duplicates will yield lower maximums and higher minimums relative to a set of independent samples, and the standard deviation will be biased low as well.

An additional key advantage of showing individual values on the nonparametric control charts that have been provided to the NMED is that the variance between duplicate samples can be easily compared to the variance between rounds. As examples, Figures 1 and 2 show nonparametric control charts for sodium and chloride, respectively from WQSP-1. The individual primary and duplicate values are shown with different symbols for each sample round. One can see at a glance that the differences between duplicates are quite similar to the differences between rounds. This indicates that any real differences between rounds cannot be accurately quantified because these differences are in the same range as the laboratory precision. Potassium concentrations from the same well, shown in Figure 3, also display some variability in duplicate results, but this variability is less than the variability between rounds. The pH data from the same well is shown in Figure 4. The agreement between duplicates for this parameter is very good, except for the two measurements in Round 7, which were anomalously low and were not reproducible. If we would have shown the average of the duplicates instead of the individual analyses, then this analytical reproducibility information would have been lost, and these trends obscured.

References

EPA, 1989, *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Interim Final Guidance*, U.S. Environmental Protection Agency, Office of Solid Waste, Waste Management Division, EPA/530/SW-89/026, February.

EPA, 1992, *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Addendum to Interim Final Guidance*, U.S. Environmental Protection Agency, Office of Solid Waste, Permits and State Programs Division, EPA/530/R-93/003, July.

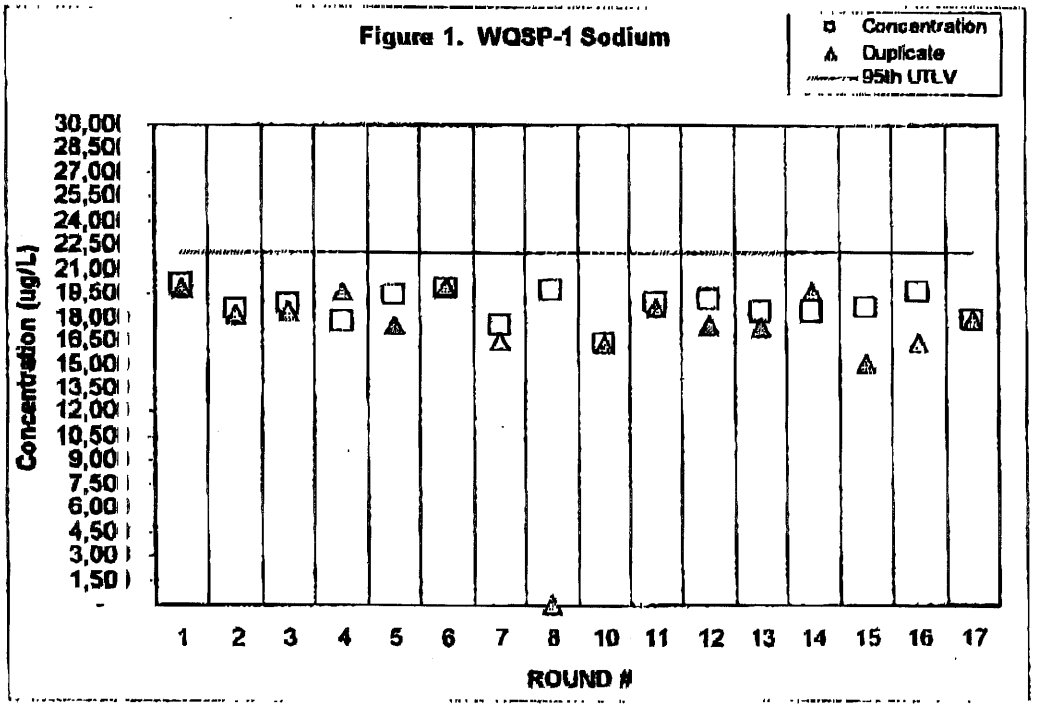
Figure 1. WQSP-1 Sodium
Concentration
Duplicate
95th UTLV
Concentration (ug/L)
30,000
28,500
27,000
25,500
24,000
22,500
21,000
19,500
18,000
16,500
15,000
13,500
12,000
10,500
9,000
7,500
6,000
4,500
3,000
1,500
1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17
ROUND #

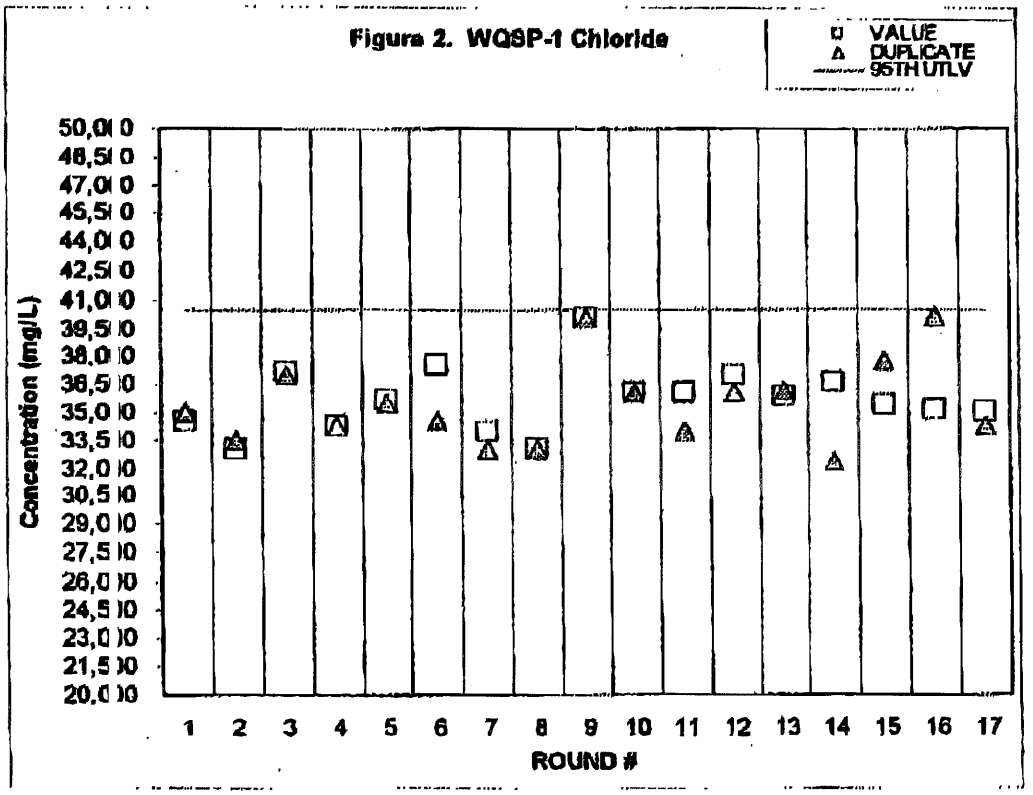
Figure 2. WQSP-1 Chloride
VALUE
DUPLICATE
95TH UTLV
Concentration (mg/L)
50,000
48,500
47,000
45,500
44,000
42,500
41,000
39,500
38,000
36,500
35,000
33,500
32,000
30,500
29,000
27,500
26,000
24,500
23,000
21,500
20,000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
ROUND #